Comprehensive Strategic Analysis of Global Subsurface Data Repositories and the Technical Framework of Routine Core Analysis Digitization

The digitization of legacy subsurface records constitutes one of the most significant challenges and opportunities in contemporary petroleum and reservoir engineering. For organizations such as NthDS, the objective is to transition from a paradigm of unstructured, scanned images to an environment of high-fidelity, machine-readable data. Routine Core Analysis (RCA) reports serve as the primary source of physical ground truth for geological formations, providing the empirical measurements of porosity, permeability, and fluid saturation necessary to calibrate secondary data sources such as wireline logs and seismic interpretations.[1] The inherent complexity of these records arises from their origin as heterogeneous documents composed of diverse elements including tabular data, semi-structured forms, technical plots, and qualitative geological summaries.[1]

The technical requirement to build ingestion pipelines that can accurately classify and extract data from these documents necessitates an exhaustive understanding of the global data landscape. This report evaluates the primary public repositories for RCA and well data, analyzes the technical architecture of legacy well files, and provides a strategic framework for creating diverse training datasets to optimize automated extraction models.

The Regulatory and Institutional Landscape of Subsurface Data

The availability of subsurface data is historically linked to the regulatory frameworks established by state, provincial, and national governments. Oil and gas operators are generally mandated to submit technical reports as a condition of their exploration and production permits. These documents, once their confidentiality period expires, become part of the public record, creating massive, high-density archives that are essential for regional geological modeling and reservoir characterization.

North Dakota Industrial Commission and the Williston Basin Archive

The North Dakota Industrial Commission (NDIC) Oil and Gas Division maintains one of the most comprehensive well-file archives in the United States, primarily focused on the prolific Williston Basin. The NDIC database is structured to facilitate chronological access to the history of a well, ranging from the initial permit application to final abandonment reports.[1] For researchers seeking RCA data, the NDIC provides several tiers of access.

The division's Premium Subscription service offers the most direct route to technical data, including scanned case files, hearing audio, digital and image well logs, and, critically, core analysis reports and core photographs.[3] The data inventory is extensive, with more than 100,600 digital and image logs available for over 18,650 wells.[3] This depth allows for the construction of training sets that span decades of shifting laboratory standards and document formatting styles. For example, a single well file like the Muller 1-21-16H (File No. 20552) can contain nearly 300 pages of data, incorporating administrative forms, directional survey reports, and laboratory-specific core analysis summaries.[1]

Data Type	**Description**	**Primary Format**	**Access Mechanism**
Well File PDFs	Comprehensive histories including RCA, DST, and completion forms	Scanned PDF	Searchable Well Index
Digital Well Logs	Gamma ray, resistivity, and porosity logs	TIFF / LAS	Subscription Portal
Core Photographs	High-resolution images of slabbed and full-diameter cores	JPEG / TIFF	Premium Subscription
Production Volumes	Monthly oil, water, and gas production statistics	Excel / HTML	Basic Subscription
Log Tops	Stratigraphic markers picked from geophysical logs	Excel	CD-ROM / Digital

The NDIC also operates the Wilson M. Laird Core & Sample Library in Grand Forks, North Dakota. This facility houses over 487,000 feet of core and an expansive inventory of drill cuttings from more than 50,000 boxes.[5] While the physical core is available for viewing and sampling by appointment, the associated analytical reports are often available in the digital well files, providing a critical link between physical rock samples and the resulting tabular data.[5]

The Texas Railroad Commission and University Lands

The Railroad Commission of Texas (RRC) is the regulatory authority for the most active petroleum provinces in North America, including the Permian Basin and the Eagle Ford Shale. The RRC's well records are a primary source for RCA data in Texas, although the documents are often categorized under specific filing forms. For instance, detailed core analyses and formation fluid sample data are frequently required for Class II injection permits and carbon sequestration projects.[6]

A significant subset of Texas subsurface data is managed by the University of Texas System's University Lands (UL), which oversees approximately 2.1 million acres of mineral interests. The UL portal provides a highly structured interface where users can search by API number to retrieve specific document categories, such as "Cores - Core Analysis" and "W-2 - Oilwell Potential Test".[7] The UL database is particularly useful for pipeline developers because it often breaks out core analysis reports as discrete files, distinct from the larger, monolithic well dossiers found in other jurisdictions.[9]

Agency	Focus Region	Key Document Types	Search Index
Texas RRC	Statewide Texas	W-2, G-1, Core Analysis Exhibits	API Number / District
University Lands	UT Mineral Interests	Cores - Core Analysis, Completion Records	API Number / Lease
Oklahoma Corp. Comm.	Statewide Oklahoma	UIC Permits, Research Lab Reports	API Number / Legal Location
Kansas Geological Survey	Statewide Kansas	Final Summary Reports, TRA Reports	Township-Range-Section

Oklahoma and Kansas State Collections

The Oklahoma Corporation Commission (OCC) holds primacy for Class II UIC regulations and maintains extensive records for wells across the state. In Oklahoma, RCA data is frequently embedded within technical research laboratory reports submitted by major operators such as Cities Service Company or EOG Resources.[10] These records often detail not just routine porosity and permeability but also specialized flow studies that are critical for understanding CO2 flood performance in sand reservoirs.[10]

In Kansas, the Kansas Geological Survey (KGS) provides highly aggregated data. Their summary reports often combine data from multiple wells within a project area, providing a broader stratigraphic context than single-well files.[11] These reports are particularly valuable for testing models that must handle multi-well correlations and complex mineralogical data, such as X-ray diffraction (XRD) and X-ray fluorescence (XRF) measurements performed on hundreds of samples.[11]

Federal Repositories and National Core Centers

For large-scale data acquisition, federal repositories offer a significant advantage by aggregating records from across state lines, particularly from federal lands and offshore territories.

The USGS Core Research Center (CRC)

The United States Geological Survey (USGS) Core Research Center in Denver, Colorado, represents the most significant publicly accessible repository of its kind in the United States.[13] It currently houses rock core from more than 9,000 wells and drill cuttings from over 52,000 wells, representing approximately 1.7 million feet of rock material from 33 states.[14]

The CRC's Well Catalog is a vital resource for subsurface researchers. It provides two primary search methods: a map-based search for geographical identification and a text-based search for comprehensive record retrieval.[15] The catalog indicates whether thin sections, core photographs, or analytical reports are available for each well.[16] If a well has been previously analyzed, the catalog often contains links to downloadable PDF files of core analysis reports, core descriptions, and stratigraphic charts.[15]

CRC Identification Type	Character Structure	Example	Description
Core Library Number	Letter + 3 Digits	A569	Unique ID for core samples

Cuttings Library Number	2 Letters + 5 Digits	CZ08825	Unique ID for cuttings samples
API Number	10-14 Digits	33-105-02157	Standard industry well identifier

The diversity of the CRC collection is exceptional, including material from standard petroleum exploration wells as well as specialized scientific drilling experiments. For instance, the collection includes data from the Cajon Pass drilling project in California, designed to study fault mechanics, and geothermal cores from Yellowstone.[17] For a data engineering project, the CRC provides an ideal source for "out-of-distribution" training data—reports from non-standard geological settings that can improve the robustness of extraction models.

Offshore and Marine Borehole Data

The Bureau of Ocean Energy Management (BOEM) and the National Oceanic and Atmospheric Administration (NOAA) manage records for the U.S. Outer Continental Shelf (OCS). These repositories focus on commercial wells and continental offshore stratigraphic test (COST) wells, providing geophysical log data and borehole records in digital formats.[19] These offshore reports are often substantially different from onshore records in terms of volume and technical focus, typically containing more detailed rock mechanics and fluid analysis data relevant to deep-water environments.[19]

International National Data Repositories (NDRs)

Organizations outside North America have frequently adopted more centralized and standardized models for subsurface data management. These National Data Repositories (NDRs) are designed to promote inward investment by making high-quality geoscientific data easily accessible to new entrants in the market.

Norway and the DISKOS Model

Norway's DISKOS system is widely acknowledged as the global gold standard for national data repositories. Established in 1992 through a collaboration between the Norwegian Petroleum Directorate and major operators like Statoil (now Equinor) and Norsk Hydro, it serves as a central hub for all seismic, well, and production data on the Norwegian Continental Shelf.[20]

The DISKOS Public Portal, managed by the Norwegian Offshore Directorate (NOD), provides a unified interface for accessing a vast array of well data types.[21] The classification system for Norwegian well data is highly structured, including:

- **Geology:** Geological reports, lithology reports, and biostratigraphy.[21]

- **Rock & Core:** Conventional core analysis (RCA), special core analysis (SCAL), and core photographs.[21]
- **Petrophysics:** Petrophysical reports and Composite Petrophysical Interpretation (CPI).[21]
- **Testing:** Fluid analysis reports, PVT analysis, and formation pressure data.[21]

The "Volve" field dataset is perhaps the most famous subset of the DISKOS archive. Released by Equinor to the public, it includes thousands of files covering every aspect of the field's lifecycle, from initial exploration to decommissioning. The Volve data has been used extensively in petrophysical research to develop models for predicting porosity and water saturation using various machine learning architectures.[22] The presence of high-quality, ground-truth RCA data paired with complete suites of wireline logs makes Norwegian data an essential component of any subsurface data development project.

United Kingdom and the North Sea Transition Authority (NSTA)

The UK National Data Repository (NDR), operated by the NSTA, mirrors the Norwegian model by providing terabytes of offshore seismic and well data.[25] A primary feature of the UK NDR is the CGG Core Analysis Database, which contains standardized core analysis data for over 2,000 wells in the UK Continental Shelf (UKCS).[26] This dataset is available for download in a massive 43GB package, providing a unique opportunity for batch processing and large-scale pipeline testing.[26]

The technical reports found in the UK NDR often focus on emerging applications, such as the Endurance and Goldeneye storage sites for carbon capture and storage (CCS). These reports provide a comprehensive technical framework for analyzing the relationship between routine porosity/permeability and the specialized flow studies required for supercritical CO2 injection.[27] UK reports also frequently include Laser Particle Size Analysis, which offers a more granular view of grain size distribution compared to traditional sieve analysis.[29]

Australia: Western Australia Petroleum Information Management System (WAPIMS)

The WAPIMS database, maintained by the Department of Mines, Petroleum and Exploration in Western Australia, is another premier source for international subsurface data. WAPIMS provides an integrated search interface that allows users to filter specifically for "CORE ANALYSIS" or "PALYNOLOGY" reports.[30]

The Australian system utilizes a "Basket" feature, enabling users to aggregate multiple reports across different wells and fields into a single download request.[30] This is particularly advantageous for creating diverse datasets from the Canning Basin, the Perth Basin, and the Carnarvon Basin, each with distinct lithologies and depositional environments.[30] WAPIMS well completion reports often include RCA as a high-density tabular appendix, requiring models to accurately identify table boundaries within documents that may exceed 500 pages.

Canada: Alberta Energy Regulator (AER) and the Oil Sands

In Canada, the Alberta Energy Regulator (AER) manages an immense archive of well data, with a specific focus on the Athabasca Oil Sands and conventional reservoirs. The AER's Product Catalogue provides access to a wide range of data, including fluid analysis (oil, water, and gas), directional surveys, and core analysis files.[32]

Alberta's reports are unique due to the focus on bitumen-bearing formations, where fluid saturation measurements (Dean-Stark extraction) are of paramount importance.[34] The AER's Digital Data Submission (DDS) system ensures that technical data is reported in consistent formats, although legacy records remain as scanned PDFs, presenting the same extraction challenges as US and UK records.[36]

Technical Decomposition of Routine Core Analysis Documents

To construct an effective extraction pipeline, one must understand the internal technical logic of RCA reports. These documents are typically generated by third-party service providers like Weatherford Laboratories, Core Lab, or Datalog and follow a structured but visually variable format.

Case Study: File No. 20552 (Muller 1-21-16H)

The Muller 1-21-16H well file from North Dakota serves as a representative example of the complexity encountered in legacy records. The file is a composite document that includes administrative, operational, and laboratory sections.[1]

1. **Administrative Precursors (Pages 1-38):** These pages contain regulatory forms, such as Form 8 (Authorization to Purchase Oil) and Form 5 (Transfer of Wells), along with legal correspondence and directional survey reports.[1] A page classification model must identify these as "Other" to prevent the extraction of irrelevant tables.[1]
2. **RCA Summary Tables (Page 39-42):** This is the high-value tabular section. The "Summary of Routine Core Analyses Results" table is structured with headers that describe the analytical conditions, such as "Vacuum Oven Dried at $180^\circ F$" and "Net Confining Stress: 2,600 psi".[1]
3. **Mineralogical Data (Page 47):** The "X-Ray Diffraction (Weight %)" table provides a detailed breakdown of clays (chlorite, kaolinite, illite), carbonates (calcite, dolomite), and other minerals (quartz, pyrite).[1]
4. **Rock Mechanics (Page 51-53):** This section contains specialized data from triaxial compressive tests, including compressive strength, static Young's Modulus, and Poisson's

Ratio.[1]

5. **Acoustic Properties (Page 52):** Tables detailing compressional and shear wave velocities (ft/sec) used to derive dynamic elastic parameters.[1]
6. **Petrographic Descriptions (Page 65-74):** Detailed thin-section descriptions that provide qualitative context for the numerical data, describing lithology (e.g., "Dolomitic, silty, very fine-grained sandstone") and texture.[1]

Key Tabular Variables and Mathematical Relationships

The primary table in an RCA report contains several critical measurements that are the focus of extraction efforts. The relationship between these variables is not just empirical but is governed by physical laws that can be used to validate the extracted data.

Parameter	Unit	Symbol	Technical Significance
Permeability	Millidarcys	k	Measures the rock's ability to transmit fluids. Often reported as k_{air} or $k_{linkenber}$.[1]
Porosity	Percent (%)	ϕ	Measures the storage capacity of the rock. Calculated as $\phi = (pv/bv) *$.[35]
Grain Density	g/cc	ρ_g	Used to infer mineralogy. Most sandstones are near 2.65 g/cc; carbonates are higher.[1]
Fluid Saturation	Percent (%)	S_w, S_o	Measures the percentage of pore space occupied by water or oil.[1]

In the Muller report, permeability is measured at a Net Confining Stress (NCS) of 2,600 psi.[1] This is a critical detail for pipeline extraction, as permeability measurements at ambient pressure are significantly different from those under reservoir stress. Furthermore, the report utilizes Klinkenberg corrections to account for gas slippage effects. The relationship between Klinkenberg permeability (k_L) and air permeability (k_a) is essential for calibrating flow models.[35]

Another important mathematical relationship found in these reports is the Formation Factor (FF), defined as:

$$FF = \frac{R_o}{R_w}$$

Where R_o is the resistivity of the rock 100% saturated with brine and R_w is the resistivity of the brine itself. The Muller report provides a "Composite m-exponent" of 1.83 derived from the Archie equation relationship $FF = \phi^{-m}$.[1] This value is essential for interpreting resistivity logs from the wellbore.

Structural Variations and Diversity in Training Sets

The user's request for diversity in input files is technically grounded in the need for model generalization. RCA reports from different eras and different labs exhibit significant structural variations.

- **Temporal Diversity:** Reports from the 1980s often feature mono-spaced typewriter fonts and simple line-drawn tables.[7] Modern reports from the 2020s utilize complex graphical layouts, embedded high-resolution photographs, and multi-layered headers.[1]
- **Service Provider Diversity:** Weatherford reports typically organize data by depth with specialized plots for core gamma.[1] In contrast, Core Lab reports may prioritize overburden vs. ambient porosity cross-plots and provide detailed "Profile Permeametry" data.[39]
- **Geological Diversity:** RCA reports for unconventional shales utilize "Tight Rock Analysis" (TRA) or "GRI" methods, which require different measurement techniques than conventional sandstone analyses.[11] These reports contain specialized variables like Total Organic Carbon (TOC) and adsorption isotherms that are absent in conventional reports.

Era	Primary Lab	Key Format Features	Extraction Challenge

1970s-80s	Bureau of Mines / Core Lab	Dot-matrix print, simple grids	OCR accuracy, paper degradation
1990s-2000s	Weatherford / Schlumberger	Digital fonts, complex appendices	Table structure identification
2010s-Present	Intertek / Stratum	High-density tables, embedded plots	Multi-modal data alignment

Algorithmic Challenges in Automated Data Extraction

The development of an automated pipeline involves overcoming several hurdles related to the unstructured nature of the source material.

Page Classification Logic

The first task is to programmatically identify "Table Pages." A successful model must distinguish between:

- **Cover Pages:** Containing logos, project numbers, and laboratory conditions.
- **Table Pages:** Characterized by high numerical density, repeating row structures, and depth-indexed entries.
- **Plot Pages:** Containing cross-plots, log charts, and core photographs.
- **Narrative Pages:** Containing thin-section descriptions or methodology summaries.

A common approach involves using a dictionary-based classifier. If a page contains keywords like "Permeability," "Porosity," and "Klinkenberg" along with a high ratio of digits to characters, it is classified as a "Table Page".[1] For example, in the provided Muller file, pages 39-42 would be classified as tables, while page 43 (a cross-plot) would be classified as "other" or "graph".[1]

Table Reconstruction and OCR Artifacts

Extracting data from "Table Pages" requires more than just optical character recognition (OCR). The pipeline must reconstruct the logical structure of the table, including:

- **Header Alignment:** Identifying which labels correspond to which columns, especially when columns are nested (e.g., "Fluid Saturations" with sub-columns "Water" and "Oil").[1]
- **Handling Sparse Data:** Many RCA tables have empty cells or cells containing special characters. In the Muller well, the "+" symbol indicates a sample unsuitable for measurement, and "**" indicates insufficient pore volume to report saturations.[1] A robust extraction engine must treat these as distinct categorical values rather than errors.

- **OCR Errors in Small Fonts:** Many laboratory reports use very small fonts for grain density or permeability values (e.g., 0.0003 md). Misinterpreting a decimal point or a leading zero can introduce significant errors into the reservoir model.

Scale and Data Integration

A second-order challenge is the integration of core data with wireline log data. Core measurements are performed on small "plugs" (typically 1 to 1.5 inches in diameter), while wireline logs average data over a volume of several cubic feet.[2] This scale difference means that extracted core data must often be "depth-shifted" to align with log features, a process that requires a high-fidelity extraction of the "Core Gamma" log usually included in the RCA report.[1]

Strategic Implementation for Bulk Acquisition

To move from a single file to a dataset of a thousand diverse RCA reports, a multi-front acquisition strategy is required.

1. **Batch Exporting from USGS CRC:** The USGS catalog allows for the export of well lists in CSV or XML.[14] By filtering for wells with "Analysis: YES," a researcher can generate a list of thousands of library numbers. These numbers can then be used to automate the retrieval of associated PDF files from the CRC's digital archive.[15]
2. **Leveraging Global NDR Databases:** The UK's CGG database (43GB) and the Norwegian DISKOS portal provide the most efficient route to high-volume, high-diversity data.[21] These systems are designed for bulk technical queries, allowing for the rapid collection of reports from varied geological basins.
3. **Utilizing Open Source Field Studies:** Datasets like the Equinor Volve release are invaluable because they provide a "complete" view of a field, allowing models to be tested against all available data types (seismic, logs, RCA, production).[22]
4. **Targeting Specific Formations for Diversity:** To ensure the pipeline can handle different lithologies, data should be collected from distinct stratigraphic units:
    - **Bakken/Three Forks (ND):** Complex, low-permeability unconventional systems.[1]
    - **Permian Basin (TX):** Prolific carbonate and sandstone reservoirs.[7]
    - **Athabasca Oil Sands (AB):** Unconsolidated sands requiring specialized cleaning.[34]
    - **North Sea (UK/Norway):** Deep-water turbidites and complex offshore completions.[41]

Synthesis and Strategic Conclusion

The technical value of Routine Core Analysis data is maximized when it is transformed from isolated, scanned documents into a cohesive, searchable, and machine-readable digital asset. Public repositories such as the NDIC, USGS Core Research Center, and international NDRs provide the raw material for this transformation. The diversity of these records—spanning different eras, laboratory protocols, and geological provinces—is essential for developing robust automated extraction pipelines.

The analysis of the Muller 1-21-16H well file demonstrates that a sophisticated ingestion model must be capable of multi-stage processing: initial page classification to isolate technical tables, followed by high-precision OCR and structural reconstruction to preserve the integrity of numerical measurements. By integrating these measurements with geomechanical data and mineralogical context, engineers can build more accurate reservoir models, leading to optimized well placement and production strategies.

Ultimately, the shift toward open subsurface data, exemplified by the DISKOS and Volve datasets, is driving a revolution in petrophysical modeling. By systematically acquiring and processing thousands of these reports, organizations can unlock legacy geological value and accelerate the development of future energy technologies, from conventional oil and gas to carbon sequestration and geothermal power.

Works cited

1. exercise-04-instructions.txt
2. Core Data | TGS Well Data Products, accessed January 29, 2026, https://www.tgs.com/well-data-products/core-data
3. Oil and Gas Subscription Services - North Dakota Department of Mineral Resources, accessed January 29, 2026, https://www.dmr.nd.gov/oilgas/subscriptionservice.asp
4. PUBLICATIONS & SERVICES, accessed January 29, 2026, https://www.dmr.nd.gov/oilgas/publications.pdf
5. Core Samples | Department of Mineral Resources, North Dakota, accessed January 29, 2026, https://www.dmr.nd.gov/dmr/ndgs/library/samples
6. RAILROAD COMMISSION OF TEXAS, accessed January 29, 2026, https://www.rrc.texas.gov/media/2s0omm1e/oxy_brp_permit_approved_10-16-25.pdf
7. Well Detail API: 42-371-34903 - University Lands, accessed January 29, 2026, https://utlands.utsystem.edu/API/4237134903
8. Well Detail API: 42-371-34690 - University Lands, accessed January 29, 2026, https://utlands.utsystem.edu/API/4237134690
9. Well Detail API: 42-003-43305 - University Lands, accessed January 29, 2026, https://universitylands.utsystem.edu/API/4200343305
10. Northeast Purdy Springer Unit (NEPSU) / South East Bradley A Unit (SEBAU) Monitoring, Reporting, a - EPA, accessed January 29, 2026, https://www.epa.gov/system/files/documents/2025-11/nepsu_sebau_mrv.pdf
11. Integrated Mid-Continent Stacked Carbon Storage Hub Project Phase II - Kansas Geological Survey, accessed January 29, 2026, https://www.kgs.ku.edu/PRS/IMSCSH/pdfs/IMSCS-HUB%20Final%20Summary%20Report.pdf
12. Site Characterization for Russell CO2 Storage ... - Regulations.gov, accessed

January 29, 2026,
https://downloads.regulations.gov/EPA-R07-OW-2025-3852-0001/attachment_15.pdf
13. Where can I get data from core samples, or access the actual cores? - USGS.gov, accessed January 29, 2026,
https://www.usgs.gov/faqs/where-can-i-get-data-core-samples-or-access-actual-cores
14. Frequently Asked Questions | U.S. Geological Survey - USGS.gov, accessed January 29, 2026,
https://www.usgs.gov/core-research-center/frequently-asked-questions
15. Resources & Services | U.S. Geological Survey - USGS.gov, accessed January 29, 2026, https://www.usgs.gov/core-research-center/resources-services
16. Well Catalog Search Tips | U.S. Geological Survey - USGS.gov, accessed January 29, 2026, https://www.usgs.gov/core-research-center/well-catalog-search-tips
17. Core Research Center | U.S. Geological Survey - USGS.gov, accessed January 29, 2026, https://www.usgs.gov/core-research-center
18. Core Handling Procedures Used for the Cajon Pass, California, Deep Drilling Experiment - USGS Publications Warehouse, accessed January 29, 2026,
https://pubs.usgs.gov/of/1989/0496/report.pdf
19. petroleum - Dataset - Catalog - Data.gov, accessed January 29, 2026,
https://catalog.data.gov/dataset/?tags=petroleum
20. Running DISKOS, Norway's data repository How Odfjell Drilling monitors platforms from shore Modelling flow in long pipelines - Finding Petroleum, accessed January 29, 2026,
http://www.findingpetroleum.com/launchmag/Issue-11/38bcdda7.aspx?r=029
21. Wells - The Norwegian Offshore Directorate - Sokkeldirektoratet, accessed January 29, 2026, https://www.sodir.no/en/diskos/wells/
22. (PDF) Analysis of Open Well Datasets - ResearchGate, accessed January 29, 2026,
https://www.researchgate.net/publication/387330114_Analysis_of_Open_Well_Datasets
23. (PDF) Building 1D and 3D static reservoir geomechanical properties models in the oil field, accessed January 29, 2026,
https://www.researchgate.net/publication/362532139_Building_1D_and_3D_static_reservoir_geomechanical_properties_models_in_the_oil_field
24. (PDF) Machine Learning Based Prediction of Porosity and Water Saturation from Varg Field Reservoir Well Logs - ResearchGate, accessed January 29, 2026,
https://www.researchgate.net/publication/361117770_Machine_Learning_Based_Prediction_of_Porosity_and_Water_Saturation_from_Varg_Field_Reservoir_Well_Logs
25. UK National Data Repository (NDR) - North Sea Transition Authority, accessed January 29, 2026,
https://www.nstauthority.co.uk/data-and-insights/data/uk-national-data-repository/
26. CGG Core Analysis Database - data.gov.uk, accessed January 29, 2026,

https://www.data.gov.uk/dataset/57fa58c2-65cc-47aa-8ef2-388a0319560b/cgg-core-analysis-database4
27. K43: Field Development Report - Technical: Storage - GOV.UK, accessed January 29, 2026, https://assets.publishing.service.gov.uk/media/5a7ffcb3e5274a2e87db72d4/K43_Field_Development_Report.pdf
28. K41: Reservoir Engineering Field Report - Category: Storage - GOV.UK, accessed January 29, 2026, https://assets.publishing.service.gov.uk/media/5a7f3fabed915d74e62294cc/K41_Reservoir_Engineering-Field_Report.pdf
29. K40: Subsurface Geoscience and Production Chemistry Reports - Technical: Storage - GOV.UK, accessed January 29, 2026, https://assets.publishing.service.gov.uk/media/5a7f8d2440f0b62305b87d51/K40_Subsurface_Geoscience_and_Production_Chemistry.pdf
30. Wells - WAPIMS | Department of Mines, Petroleum and Exploration, accessed January 29, 2026, https://wapims.dmp.wa.gov.au/WAPIMS/Search/Wells
31. Geological Survey and Resource Strategy Annual Review 2019–20, accessed January 29, 2026, https://warmelpdstageodocspub.blob.core.windows.net/gswa-publications/0fa88a98-b7c5-4ff1-911d-3afaab3554d7.pdf
32. AER Products and Services Catalogue – Wells - Alberta Energy Regulator, accessed January 29, 2026, https://www1.aer.ca/productcatalogue/WELL.html
33. AER Products and Services Catalogue – Home - Alberta Energy Regulator, accessed January 29, 2026, https://www1.aer.ca/productcatalogue/index.html
34. AER/AGS Report 101: Revised Bedrock Topography and Characterization of Quaternary Sediments in the Fort McMurray Region, Northea - Alberta Geological Survey, accessed January 29, 2026, https://ags.aer.ca/document/REP/REP_101.pdf
35. Routine Core Analysis - datalog, accessed January 29, 2026, http://www.datalogme.com/index.php/services/core-analysis/routine-core-analysis
36. Directive 013 and Inactive Well Compliance Program FAQs - Alberta Energy Regulator, accessed January 29, 2026, https://www.aer.ca/documents/directives/Directive013andIWCP-FAQs.pdf
37. Disclaimer This Report, including the data and information contained in this Report, is provided to you on an "as is" and t - Open Government program, accessed January 29, 2026, https://open.alberta.ca/dataset/2ef6c060-69a5-4bde-9ae5-370f6b9219fc/resource/96e689ef-b8b1-4fca-9f54-3b9330ed8c9b/download/questmmvplan_2016.pdf
38. Petrophysical Characterization of Reservoirs in the Pil and Bue area,Norwegian Sea, accessed January 29, 2026, https://nva.sikt.no/registration/0198e882a0af-263fe5b2-af50-45b9-905a-08b52ec7da5c
39. Routine Core Analysis Well Thylacine #1 Offshore Otway Basin Australia, accessed January 29, 2026,

https://www.mrt.tas.gov.au/mrtdoc/petxplor/download/OR_0558/THYLACINE_1.pdf
40. (PDF) The Design of an Open-Source Carbonate Reservoir Model - ResearchGate, accessed January 29, 2026, https://www.researchgate.net/publication/358582459_The_Design_of_an_Open-Source_Carbonate_Reservoir_Model
41. PVT Modelling Report for CO2 in Goldeneye Project - GOV.UK, accessed January 29, 2026, https://assets.publishing.service.gov.uk/media/5a7f8c4340f0b6230269026d/Petrophysical_Modelling_Report.pdf